

2017-03-30

In Defence of Effect-Centric Research

Wills, Andy

<http://hdl.handle.net/10026.1/9181>

10.1016/j.jarmac.2016.10.005

Journal of Applied Research in Memory and Cognition

American Psychological Association (APA)

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

In defence of effect-centric research

Andy J. Wills & Timothy J. Hollins

Cognition Institute, School of Psychology, Plymouth University

For correspondence regarding this manuscript, please contact

Prof Tim Hollins, School of Psychology, Plymouth University, Plymouth. PL4 8AA,
U.K, or email thollins@plymouth.ac.uk.

This is the author's accepted manuscript. The final published version of this work (the version of record) is published in *Journal of Applied Research in Memory and Cognition*, 6, 43-46.

10.1016/j.jarmac.2016.10.005. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

Authors' notes

The authors contributed equally to this commentary.

In their target article, De Houwer, Hughes and Barnes-Holmes describe an archipelago of islands each colonised by a different tribe of psychologists, with little awareness of their neighbours. They focus particular attention on two tribes – the cognitivists and the functionalists – who appear to share many interests (they study the same topics, try to solve the same problems), but who have diverged in their scientific methods and theories. Worse, these two tribes communicate very little, and so don't learn from each other. The authors are troubled by this, and so offer the tribes a solution, focusing on application.

The paper's central claims are that, 1) the tribes different approaches should not be seen as being in conflict, but as mutually supportive, 2) progress in understanding topics (or solving problems) would be enhanced if the tribes communicated, 3) different ways of combining functional and cognitive approaches are possible, 4) the focus of the communication should be on the shared functionalist core, 5) most progress will be made by taking an analytic-abstractive theoretical approach and 6) the applied members of the two tribes could merge to become psychological engineers. Broadly, our agreement is negatively correlated with position in this numbered list. Consequently, we are doubtful that the specifics of their proposal will achieve their intended goal ... And we really don't want to be called psychological engineers.

Shared perspectives

Largely, we agreed with the thrust of the early sections of the paper. Different functional (behaviourist) and cognitive (mentalist) accounts of a behavioural phenomenon should be possible without the two accounts being antagonistic. Communication between areas is both valuable and to be encouraged. Indeed, we

would go further and argue that there are yet other approaches carried out by other tribes on other islands who could usefully join the conversation. The functional imaging tribe have been particularly active of late, but there are other tribes (biologists, pharmacologists, sociologists, anthropologists...) who each have something valuable to say about diverse behavioural phenomena.

We further agree that a functional approach to describing phenomena is central to good psychological science, irrespective of one's theoretical position. It is important to have a clear understanding of the relationship between independent and dependent variables; such is the knowledge base upon which theories are inspired and tested. In essence, this is simply good scientific practice, because a good experimental design entails the precise specifications of the environmental conditions required to bring about a behavioural effect. The failure to maintain high standards in this regard is part of the underlying cause of the recent replicability crisis in science. Unless one is very sure of a phenomenon, generating explanations (i.e. taking a theoretical position) is premature. Thus, to us, method and findings trump theorising.

We also agree that a pervasive problem in cognitive psychology is to confuse empirical phenomena with the theories that seek to explain why those phenomena occur. De Houwer et al.'s term 'proxies' provides a useful label for this error. Without wishing to raise heckles by picking upon specific examples, it strikes us that broad concepts such as “executive function”, “inhibition” or “metacognition” are often used loosely, and in ways not justified by the highly paradigm specific measures employed. As an illustration of this danger, Shilling, Chetwynd and Rabbitt (2002) looked at age-related changes in “inhibitory deficits” across 4 variations of the classic Stroop task involving stimuli that could be processed in two conflicting ways (based

on colours, numbers, spatial location and figure vs ground conflicts). All four tasks showed the predicted age-related pattern, but the shared variance between tasks was zero. That is, there was no overall inhibitory deficit that the measures shared, despite the surface similarity in the tasks.

Finally, we agree with what we take to be one of the main intents of the target article - to encourage and develop a common language - by which researchers of different theoretical backgrounds can communicate and collaborate effectively. However, where we start to diverge from the authors is whether a *new* shared language is required.

A parting of the ways

De Houwer and colleagues argue that “*communication between applied psychologists can be improved by adopting a common analytic-abstractive language that is phrased in terms of general functional principles*” (p.28). They go to some length to explain that this functional analytic-abstractive language is a language of explanations (theories), rather than a description of effects (which they call “functional effect-centric” research). Functional analytic-abstractive theories, they argue, are expressed in sufficiently general terms that they can make predictions, and those predictions are testable (and hence potentially wrong).

To us, this proposal risks substituting one kind of proxy error with another. Just as cognitive psychologists often confuse their explanations with observations so too, it seems, functional psychologists sometimes confuse their non-cognitive abstractive theories with the observations they are intended to explain.

For example, we have no problem with the authors' assertion that “both the lever pressing of a rat in a Skinner box as well as the tantrum of a small child at home can be described as instances of the general functional principle of reinforcement”, nor with their assertion that such a theory allows one to theorize that one could “reduce the frequency of tantrums by instructing a parent to ignore the child when it has a tantrum”. These are indeed good examples of a theory couched in functional terms. However, at this level of abstraction, a functional account is no more a description of specific observed phenomena than is a theory expressed in cognitive terms. The authors' proposal – that psychologists should talk in a common language of functional abstractions – seems unlikely to lead to progress. Scientists of different scientific persuasions can profitably communicate in terms of observed phenomena; communicating in terms of different abstractions is likely to be less productive.

To illustrate this point, consider heavenly bodies. From a common set of observations, Copernican astronomers proposed that planets moved in perfectly circular orbits, whilst Keplerian astronomers proposed elliptical orbits. How should these two astronomical tribes talk to each other? De Houwer et al.'s proposal is akin to the suggestion that they should communicate in the common language of, say, Copernican circular orbits. However, Keplerian astronomy turned out to be the better theory. It seems there would have been little to be gained in encouraging Keplerian theorists to use the common language of Copernicus. The common language of science, we argue, should be the observations, both those already made, and those one would wish to make in the future to distinguish the theories. Of course, Keplerian astronomers did eventually convince their field to theorize in their terms, but this was because their theory provided a better explanation of the observed phenomena.

Perhaps, one day, either functional or cognitive theories will 'win' in a similarly decisive manner --- but that day has not yet arrived.

We believe that the common language for functional and cognitive theorists (and the other tribes) should be known observations and (possibly diverging) predictions. There seems to be some chance that people of different theoretical persuasions might agree on what has been reliably observed, and so what is worth testing in the future (perhaps because they disagree on what will be observed when such observations are made). Such an approach is increasingly used to establish common ground between different theoretical approaches within cognition, such as adversarial collaborations in which the predictions are agreed in advance (e.g. Dwyer & Waldmann, 2016) or distributed collaboration in which data sets that are independently replicated are shared to facilitate formal model testing (e.g. Wills et al., 2017, Wills & Pothos, 2012). Thus we think that what De Houwer et al describe as 'effect-centric functional research' provides the ideal basis, and indeed probably the only workable basis, for collaboration between researchers of different theoretical persuasions and moreover that formal means for communication and collaboration already exist. We also believe that it provides the best approach to applied research, as we outline below.

Another of De Houwer et al.'s proposals is that the "cognitively-inspired analytic-abstractive approach to applied psychology combines the best of both worlds (i.e. levels of explanation) and can thus maximize progress" (p. 27, emphasis added). We are skeptical about this proposal, and the example given in the target article concerning dot-probe and Stroop effects (p.25--27) did not convince us. Their example begins (on page 25) with a purely functional account of the dot-probe task. They then go on to say:

“... functional knowledge about proportion-congruency effects in Stroop studies *and cognitive models of these effects* can be used to inspire research on dot-probe-based training. For instance, it has been demonstrated that proportion-congruency effects are highly stimulus-dependent (e.g., Schmidt & Besner, 2008) *suggesting that these effects are not due to conflict adaptation via changes in the deployment of attention to task-irrelevant stimuli. Instead, proportion-congruency effects have been explained by episodic memory models that operate solely on the basis of general principles of memory storage and retrieval (e.g., Schmidt, 2013).* This functional *and cognitive* knowledge about proportion-congruency effects in Stroop tasks sheds new light on the fact that dot-probe-based training effects show little transfer (i.e., also seem to be stimulus-specific; see Koster & Bernstein, 2015, for a review). (p 26-27).

It seems to us that the (rather minimal) cognitive theorizing in this example could be removed without much being lost. As an exercise, re-read the quotation above, skipping over the italicised text. To us, this does not affect the argument, which appears to be made on purely functional principles. We found the functional argumentation rather interesting, and a good illustration of the utility of a functional analytic-abstractive approach. It's a similar feeling we get when reading a good piece of cognitively-inspired prediction. It just seems to us that it would be better, and more realistic, to assume that theorists will use whatever theoretical approach they are most familiar or content with, and count on the presumption that for different researchers, this will be different theories. Thus, the field as whole can benefit from a plurality of approaches, without assuming such plurality need reside within individual researchers, or that all researchers should adopt such plurality.

As an aside, we think it rather telling that the authors acknowledge that the kind of research they believe ought to be popular (cognitively-inspired analytic-abstractive)

actually isn't, and that "there are few if any examples of this type of research in the applied psychology literature" (p.17). They attribute this absence to the troubled history between the cognitive and functional tribes, but we suspect that if it were a valuable way forward, enterprising researchers would have discovered it by now. Moreover, given that the authors themselves struggle to generate a convincing example of such an approach, we suspect that the approach may not be particularly fruitful, and are therefore unconvinced that it represents "a promising avenue for applied researchers" (p.17).

We also question what is meant by "maximising progress" in applied research. If the goal is draw out general principles, then we accept that broader levels of analysis and theorising may be appropriate. However, the broader the theorising, then potentially the less precise the prediction that an explanation (theory) provides in any particular situation, and so potentially the lower the utility and testability of the account. Too many psychological theories are cast in this way, and so very hard to apply with any confidence to specific situations. For instance, there is a wealth of research looking at the effects of stress on memory, but this is so vaguely cast that it is impossible to predict with any certainty how a particular individual will recall a particular stressful event, and at the same time, impossible to assert with any authority that a particular observation refutes the theory.

In contrast, if the goal is to optimise a behavioural outcome in a particular setting, then we believe that the most progress is likely to be made by designing experiments that mimic that situation as closely as possible. That is, to conduct well motivated effect-centric research. For example, in the field of eyewitness identification research the goal is to discover ways of conducting identification tests that maximise the quality of evidence obtained. Consequently, researchers strive to mimic key

aspects of the identification task, and then manipulate it in ways that are either purely functional (e.g. Levi, 1998), or cognitively-inspired (Perfect & Weber, 2012). In either case, the goal is the same, and is specific to lineup choices: there is no aim to draw abstract conclusions about the nature of decision making, or to generalise to other kinds of tasks.

It is also possible to think of examples where different researchers address the same topic with different orientations. Recent research on the testing effect can be cast in terms of the theoretical understanding of the cognitive (or functional) principles illuminated by the testing effect (e.g. Knight, Ball, Brewer, DeWitt & Marsh, 2012), or in terms of an attempt to maximise student performance on a final test (e.g. Pan, Gopal & Rickard, 2016). Researchers with the first orientation are likely to take an abstractive theoretical approach and seek to understand the basic principles that might apply whenever material is reprocessed after a delay. Those with a more applied approach are interested to know the conditions (the delay, the materials, the instructions etc.) under which testing is beneficial. It is hard to generate an approach that combines these two approaches that will maximise progress, because they are progressing toward different goals.

Our central point here is that in order to maximise progress, one must know what one is progressing towards, and we suspect that the majority of applied researchers would want to make progress on specific problems. In part, we suspect that this reflects the relatively poor evidence for far-transfer effects in psychology. To take a single recent example, the evidence strongly suggests that “brain training” has large effects on the trained skill, but small to non-existent benefits on distantly related tasks or on everyday life (Simons, Boot, Charness et al, 2016). The corollary of this point is that if one wishes to solve real-world problems, then progress is best made

by creating laboratory analogs that are as close to those problems as possible: abstracting away from those problems may lead to failures of transfer. Thus, there seem to be some dangers inherent in an overly-abstractive approach, particularly in applied contexts.

Should JARMAC become the Journal for Psychological Engineering?

Finally, we'd caution against adoption of the term 'psychological engineering' for both sociological and conceptual reasons. From the perspective of 20th century history and dystopian science fiction, we find the term 'psychological engineering' vaguely terrifying. We're not sure we want our psyche engineered, and we suspect the term would win applied psychologists few friends in the wider media. The authors must, we are sure, be aware of the surface-similar term 'social engineering' and its use within information security to refer to making use of peoples' cognitive and social weaknesses to obtain confidential information. However effective such approaches are (and they do seem rather effective), it is presumably not a sphere of human endeavour with which most applied psychologists would want to associate.

At a deeper level, the term 'psychological engineering' will inevitably be contrasted with the pre-existing term 'psychological science'. We doubt that many applied psychologists would consider the relationship between applied and non-applied psychology to be similar to the relationship between science and engineering. The flow of theoretical knowledge between science and engineering is largely one way. For example, civil engineering applies known theory in physics to build bridges that don't (usually) fall down. The theory used is Newtonian, and the theory has remained unchanged for centuries. Newtonian theory is unlikely to be further refined by civil engineers, however many more bridges they build. Nor is it likely that, in the absence

of the advances in theoretical physics through the 20th Century, Newton's laws would have been overturned by engineers.

In contrast, we suspect many applied psychologists, particularly those who consider themselves to be 'cognitive', are motivated, at least in part, because of the potential to contribute to psychological theory. At least one major applied journal (the *Journal of Experimental Psychology: Applied*) describes its explicit aim as being to bridge practical problems and psychological theory. We suspect some applied researchers believe that working on applied problems is sometimes a better way to advance theory over entirely non-applied laboratory work. They may be right, although curiosity-based science has also had its share of successes.

Author contributions

Both authors contributed equally to the preparation of this manuscript. We shared ideas, authorship, proof-reading-duties and several coffees during its preparation.

References

- Dwyer, D. M. & Waldmann, M. R., (2016) Beyond the information (not) given: representations of stimulus absence in Rats (*Rattus Norvegicus*). *Journal of Comparative Psychology*, 130, 192-204.
- Knight, J. B., Ball, H. B., Brewer, G. A., DeWitt, M. R. & Marsh, R. L. (2012). Testing unsuccessfully: a specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language*, 66, 731-746.
- Levi, A. (1998). Protecting innocent defendants, nailing the guilty: a modified sequential lineup. *Applied Cognitive Psychology*, 12, 265-275.
- Pan, S. C., Gopal, A., & Rickard, T. C. (2016). Testing with feedback yields potent, but piecewise, learning of history and biology facts. *Journal of Educational Psychology*, 108, 563-575.
- Perfect, T. J. & Weber, N. (2012). How should witnesses regulate the accuracy of their identification decisions: one step forwards, two steps back? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 38, 1810-1818.
- Shilling, V. M., Chetwynd, A. & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: an investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, 40, 605-609.
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z. & Stine-Morrow, E. A. L. (2016). Do “brain-training” programs work? *Psychological Science in the Public Interest*, 17, 103-186.

Wills, A.J., O'Connell, G., Edmunds, C.E.R., & Inkster, A.B. (2017). Progress in modeling through distributed collaboration: Concepts, tools, and category-learning examples. *Psychology of Learning and Motivation*, *accepted*.

Wills, A. & Pothos, E. M. (2012). On the adequacy of current empirical evaluations of formal models of categorization. *Psychological Bulletin*, *138*, 102-125